

WSNEWS Federated OpenSearch Server and Client

ESIP Federated Search Workshop
Tuesday, July 20, 2010

Hook Hua, Eric Fetzer, Matthew Henderson, Steven Lewis
NASA/Jet Propulsion Laboratory

Copyright 2010 California Institute of Technology.
Government sponsorship acknowledged.

The Need

SIMPLE DATA SHARING

Sharing Data

- Many Projects with Data.
 - Many tiers
- Data search and access. *“Why so difficult?”*
- Federated data holdings



NEWS Global Water and Energy Budget Integration Data Sets

SRB-RFA/GEWEX	surface shortwave, longwave, and net radiation	global	7/1983-12/2004	2.5 deg X 2.5 deg	monthly	NASA Langley DAAC; http://eosweb.larc.nasa.gov	satellite	Bing Lin (bing.lin@nasa.gov)	Lin	?	Energy Cycle	long-term
SRB	surface shortwave, longwave, and net radiation	global	7/1983-12/2004	equal-area (1 deg X 1 deg at the equator)	3-hourly, daily and monthly	NASA Langley DAAC; http://eosweb.larc.nasa.gov	satellite	Bing Lin (bing.lin@nasa.gov)	Lin	?	Energy Cycle	long-term
ISCCP-RFA/GEWEX	top of atmosphere and surface shortwave, longwave, and net radiation	global	7/1983-12/2004	2.5 deg X 2.5 deg	monthly	NASA Langley DAAC; http://eosweb.larc.nasa.gov	satellite	Bing Lin (bing.lin@nasa.gov)	Lin	?	Energy Cycle	long-term
ISCCP	top of atmosphere and surface shortwave, longwave, and net radiation	global	7/1983-12/2004	pixel and gridded	instantaneous, daily, and monthly	GISS; http://isccp.giss.nasa.gov	satellite	wrossow@qiss.nasa.gov	Lin	?	Energy Cycle	long-term
SeaFlux (GSSTF-2)	latent and sensible heat fluxes at the sea surface	global	1989-2000	1 deg X 1 deg	daily	NASA Goddard (http://daac.gsfc.nasa.gov/precipitation/gsstf2.0.shtml)	satellite	curryja@eas.gatech.edu	Lin	? & ?	Energy Cycle	long-term
SeaFlux (HOAPS)	latent and sensible heat fluxes at the sea surface	global	1989-2000	1 deg X 1 deg	daily	Meteorological Institute of Univ. Hamburg; http://www.hoaps.zamg.at	satellite	curryja@eas.gatech.edu	Lin	? & ?	Energy Cycle	long-term
Global Land Data Assimilation System (GLDAS)	surface net and downward shortwave and longwave; latent, ground, and sensible heat fluxes; soil temperature (levels); others	90N to 60S	1/1979-present	1 deg and 0.25 deg	3-hourly	NASA Goddard; http://ldas.gsfc.nasa.gov	land surface model data integration	Matthew.Rodell@noaa.gov	Lin	?	Energy Cycle	long-term; model assimilation
JPL ocean study (Tim Liu)	ocean surface momentum (wind-stress) and fresh water flux (including latent heat)	60N-60S	8/99 to present			JPL; http://airsea-www.jpl.nasa.gov/	satellite	liu@pacific.jpl.nasa.gov	Lin	?	Energy Cycle	long-term
sea-surface latent heat flux (Frank Wentz)	latent heat flux at the sea surface	global ocean	July 1987 - present	0.25 deg x 0.25 deg	monthly	Remote Sensing Systems (www.remss.com)	satellite + in situ	support@remss.com	Lin	?	Energy Cycle	long-term; combined with other RSS data
Ocean Wind and Moisture (Frank Wentz)	sea surface temperature, surface wind speed and direction, columnar water vapor, columnar cloud water, surface rain rate	global ocean	July 1987 - present	0.25 deg x 0.25 deg	daily, 3-day, weekly, monthly	Remote Sensing Systems (www.remss.com)	satellite: SSM/I, TMI, AMSR-E, QuikSCAT	support@remss.com	Lin	?	Energy Cycle	long-term; combined with other RSS data
latent heat profile	profiles of atmospheric heating due to phase changes of water substance; eddy transports	37N to 37S; oceans	10/1998 - 3/1999; 01/2001 - 07/2001	0.5 deg X 0.5 deg	instantaneous	NASA GSFC	satellite	olson@agnes.gsfc.nasa.gov	Lin	No	Energy Cycle	short-term
radiative heating profile	vertical profiles of LW, SW, and net radiation and radiative heating	40N to 40S	1/1998-12/2001 (additional months being processed)	0.5 deg X 0.5 deg X 1 km	instantaneous	Colorado State University	satellite	tristan@atmos.colostate.edu	Lin	No	Energy Cycle	short-term
Precip. (latent heat; R. Adler; W. Tao)	GPCP global surface precip.	global	1/1979-9/2006	1 deg X 1 deg	daily	NASA Goddard; http://precip.gsfc.nasa.gov	satellite	Robert.F.Adler@noaa.gov	Lin	?	Energy Cycle	long-term

Desirements of OpenSearch Server

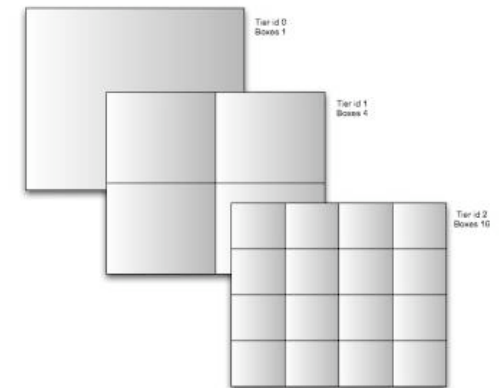
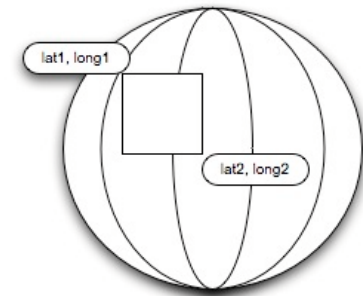
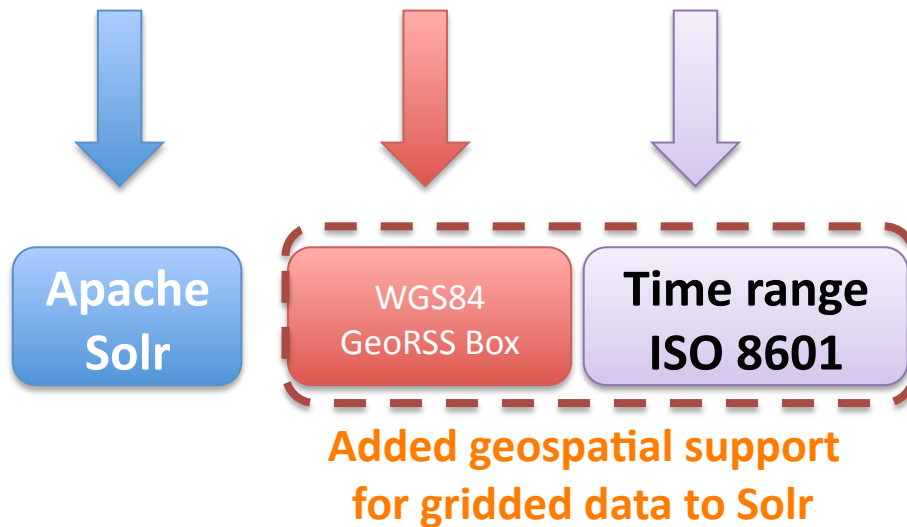
- **Portable**
 - Easy relocation of services
- **Light-weight**
 - Small server foot print
 - No large dependencies
- **Simple to deploy**
 - “Untar and run” with default settings
 - Little or no setup
 - No root privileges needed

Federated OpenSearch

SERVER

OpenSearch Components

- Full-text + space + time requests

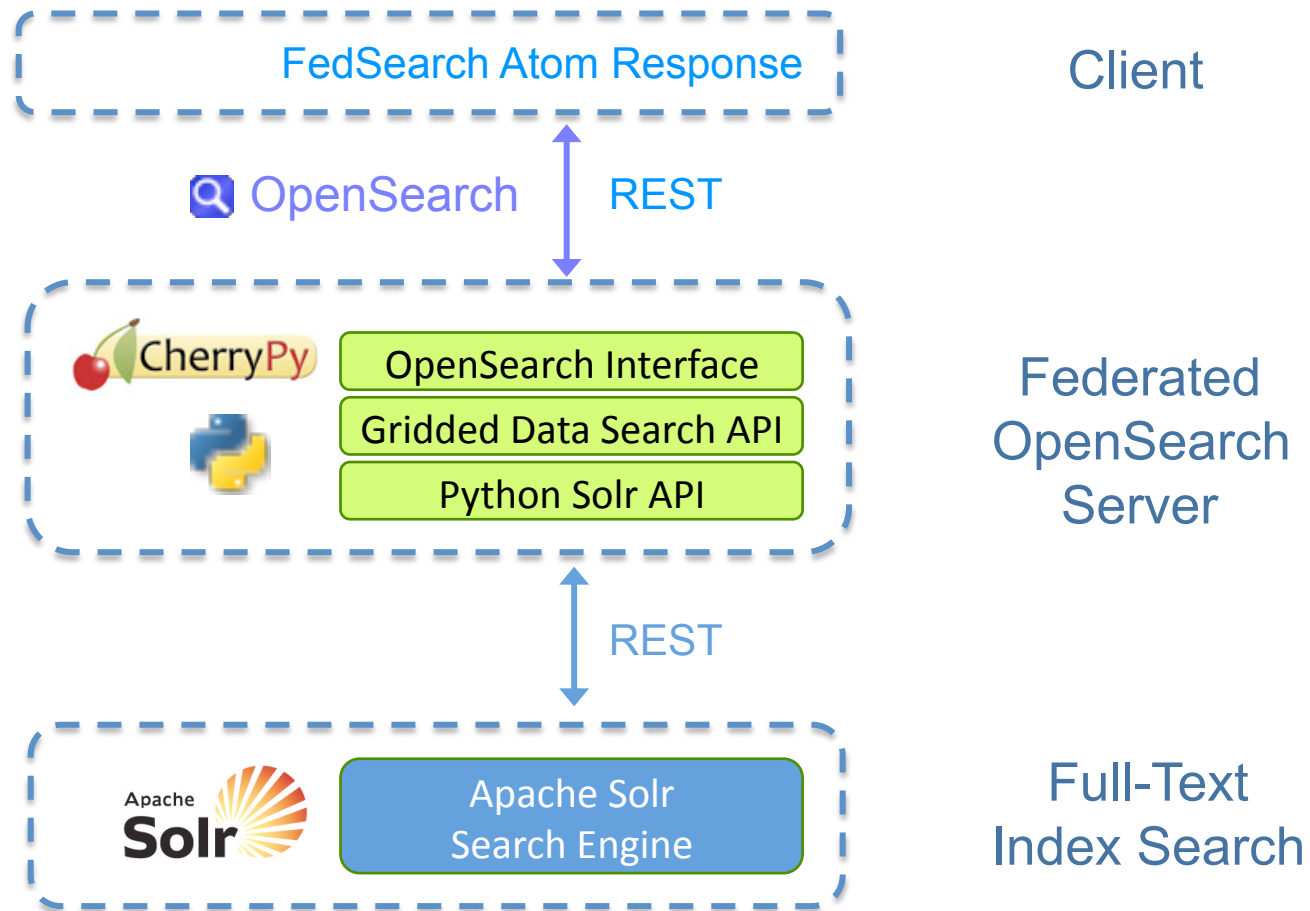


- Local Lucene
- Solr SpatialSearch
- Issue tracking SOLR-773: incorporating Local Lucene into Solr

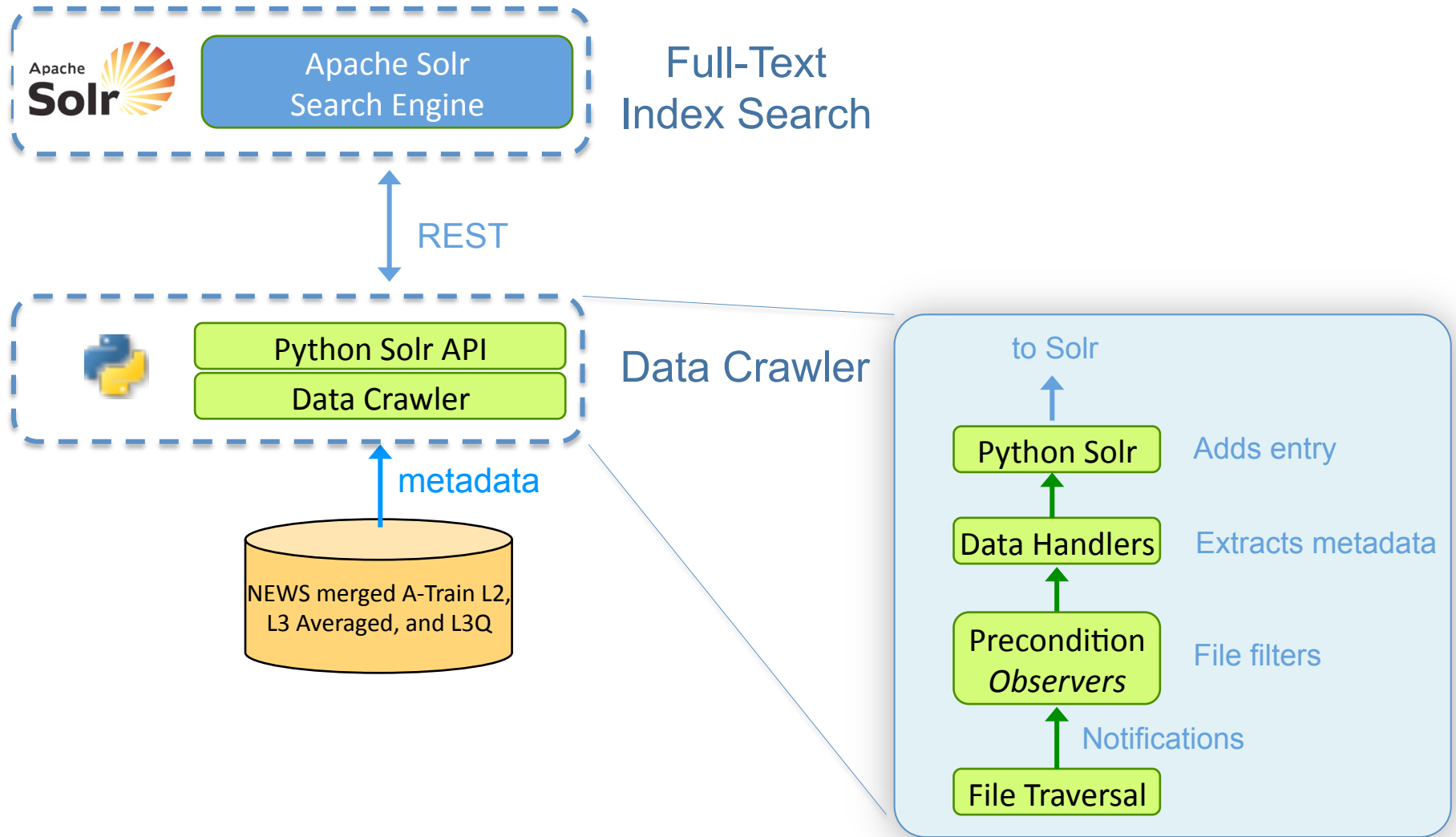
Apache Solr

- Open source enterprise search platform
- Leverages Apache Lucene for
 - full-text indexing and search
 - Rank/scoring of results
 - REST-like HTTP/XML and JSON APIs
- Faceted Search
- Pagination
- Distributed Search – *sharding*
- Handles ~100 million documents per single machine
- Multi-core support

Search Layers



Getting Data In





Federated OpenSearch Response

- Leveraged Google GData API
 - Google Data Protocol is a REST-inspired technology for reading, writing, and modifying information on the web.
- Atom <content> element used for user-friendly info with *type=html*.
- Extended Atom response
 - WGS84 GeoRSS Box
 - ISO 8601 date time representation
 - rel=<http://esipfed.org/ns/fedsearch/1.0/namespaces>

Example Response

```
<ns0:feed xmlns:ns0="http://www.w3.org/2005/Atom" xmlns:geo="http://a9.com/-/opensearch/extensions/geo/1.0/" xmlns:time="http://a9.com/-/opensearch/extensions/time/1.0/">
  <ns0:title>temperature - Dataset "NEWS Merged 3-hour Average 1-degree" - WSNEWS OpenSearch Granules</ns0:title>
  <ns0:subtitle>Showing results 0 to 9 (10 of 3409 found) for "temperature", bounding box: (-180.0, -90.0, 180.0, 90.0), datetime range: [2000-01-1T00:00:00Z to 2010-01-1T00:00:00Z]. Query time: 1 ms </ns0:subtitle>
  <ns0:author><ns0:name>WSNEWS/JPL</ns0:name></ns0:author>
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100" />
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100/opensearch" rel="http://esipfed.org/ns/fedsearch/1.0/search#" />
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100/osdd_granule?dataset=NEWS Merged 3-hour Average 1-degree" rel="search" type="application/opensearchdescription+xml" />
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100/opensearch/granule?dataset=NEWS Merged 3-hour Average 1-degree&q=temperature&bbox=-180,-90,180,90&datetimestart=2000-01-1T00:00:00Z&datetimeend=2010-01-1T00:00:00Z&p;startindex=0&itemsperpage=10&format=atom" rel="self" />
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100/opensearch/granule?dataset=NEWS Merged 3-hour Average 1-degree&q=temperature&bbox=-180,-90,180,90&datetimestart=2000-01-1T00:00:00Z&datetimeend=2010-01-1T00:00:00Z&p;startindex=0&itemsperpage=10&format=atom" rel="first" />
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100/opensearch/granule?dataset=NEWS Merged 3-hour Average 1-degree&q=temperature&bbox=-180,-90,180,90&datetimestart=2000-01-1T00:00:00Z&datetimeend=2010-01-1T00:00:00Z&p;startindex=10&itemsperpage=10&format=atom" rel="next" />
  <ns0:link href="http://wsnews.jpl.nasa.gov:8100/opensearch/granule?dataset=NEWS Merged 3-hour Average 1-degree&q=temperature&bbox=-180,-90,180,90&datetimestart=2000-01-1T00:00:00Z&datetimeend=2010-01-1T00:00:00Z&p;startindex=3399&itemsperpage=10&format=atom" rel="last" />
  <ns0:updated>2010-07-20T00:00:29Z</ns0:updated>
  <ns0:entry>
    <ns0:content type="html">
      ...
    </ns0:content>
    <ns0:updated>2009-04-22T11:38:20Z</ns0:updated>
    <ns0:title>2002_06_01_blk_7.nc</ns0:title>
    <ns0:link href="http://wsnews.jpl.nasa.gov:8100/file/download_by_id?id=48bc660bc494bb27dc86820faf089a1e35ae6f57" length="760584" title="Granule File" type="application/x-netcdf" />
    <ns0:link href="http://wsnews.jpl.nasa.gov:8100/file/download_by_id?id=48bc660bc494bb27dc86820faf089a1e35ae6f57" length="760584" rel="http://esipfed.org/ns/fedsearch/1.0/data#" title="Granule File" type="application/x-netcdf" />
    <geo:box>-180.0, -90.0, 180.0, 90.0</geo:box>
    <time:start>2002-06-01T18:00:00Z</time:start>
    <time:end>2002-06-01T21:00:00Z</time:end>
  </ns0:entry>
</ns0:feed>
```

WSNEWS OpenSearch Service

- Currently about 55K granules available

- Level 2 Regrid
- Level 3 Averages
- Level 3Q Summaries

- OSDDs

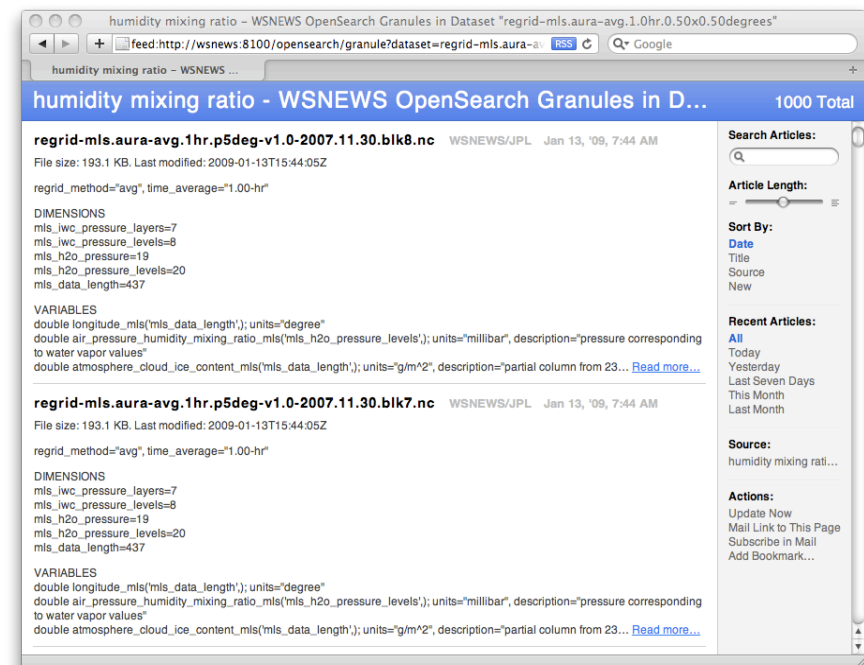
- Full-text + space + time

- Dataset-level search
- Granule-level search

- Direct downloads

- Linked with Mirador (GSFC)

Search Results as News Cast Feed



<http://wsnews.jpl.nasa.gov:8100/opensearch/>

Federated OpenSearch

DOWNLOADER CLIENT

Use Case

- *“Search for and download all granules in a dataset”*
- Atom readers for OpenSearch insufficient
- Need automated downloads of search results
- Keep it simple!
 - So simple why wouldn't you use it

opensearchget

- Python API for Federated OpenSearch “standard”
- Command-line tool
- Make it simple to autodownload search results

```
Usage: opensearchget.sh [options] "<dataset>" "<keywords>"
```

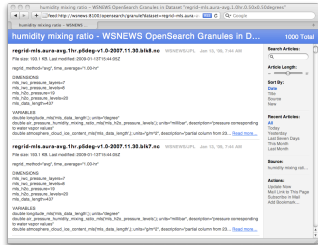
```
opensearchget.sh \"NEWS Merged 3-hour Average 1-degree\" \"temperature\"
```

```
opensearchget.sh --paginateStart=0 --paginateRows=10  
\"NEWS Regrid amsre.aqua 1.0hr Average 0.50x0.50degrees v1.0\"  
\"temperature\"
```

```
opensearchget.sh --bbox=-180,-90,180,90  
--datetimeStart=2008-01-01T00:00:00Z --datetimeEnd=2008-12-31T23:59:59Z  
--paginateStart=0 --paginateRows=10  
\"NEWS 1.0month Averages 0.50x0.50degrees v1\" \"temperature\"
```


Multiple Ways to Access

1



WSNEWS at JPL

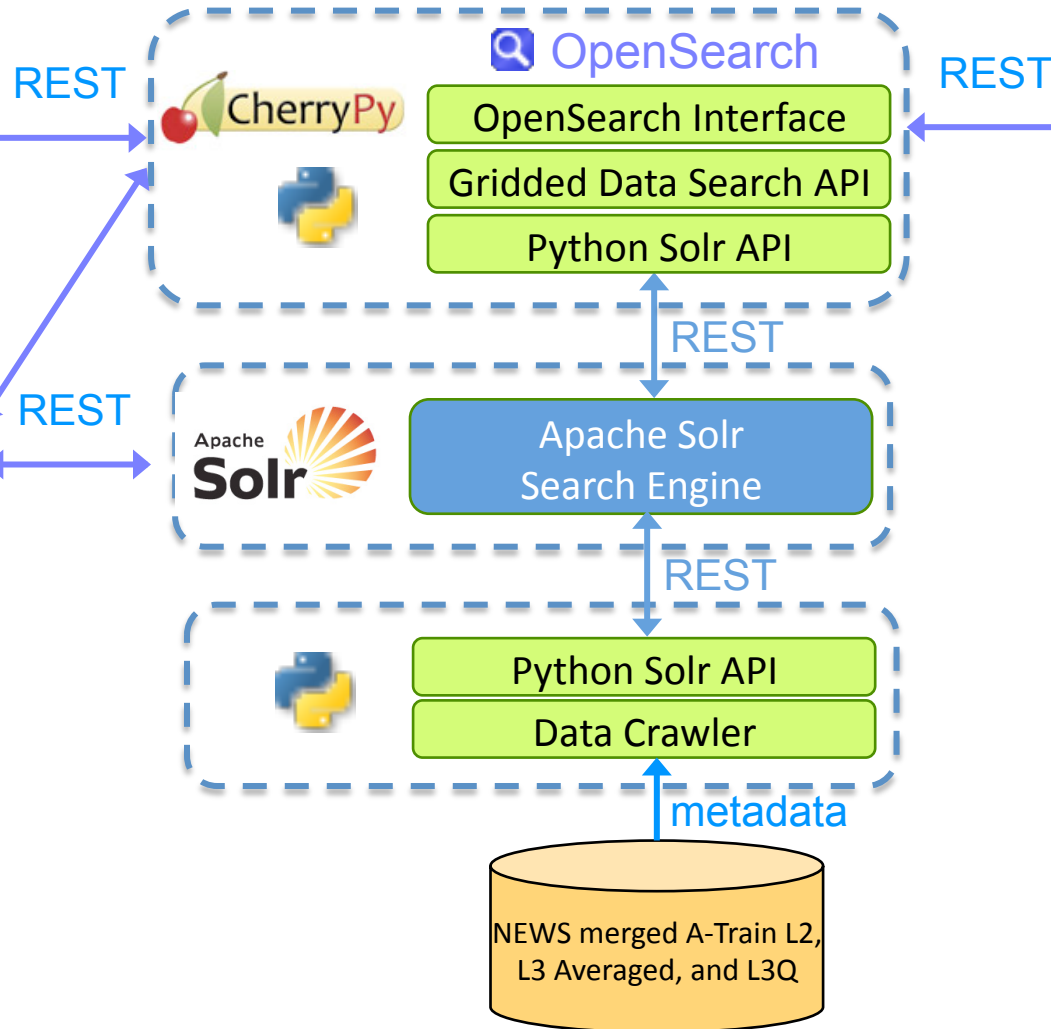
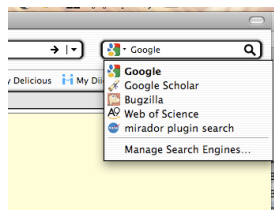
Mirador at GSFC

2

Command-line

3

Browser search bar



Water and Energy Cycles	Downward Shortwave Flux	Surface Shortwave Flux	Upward Longwave Flux
Atmospheric Radiation The process by which electromagnetic radiation is propagated through free space.	Project: MERRA VARS	Cloud Effective Radius	Cloud Fraction
Clouds A visible aggregate of minute water droplets and/or ice crystals in the Earth's atmosphere.	Cloud Condensation Nuclei	Cloud Liquid Water	Cloud Optical Thickness
Heat Flux Heat flux is the amount of heat that is transferred across a surface of unit area in a unit of time. Also refers to latent and sensible heat fluxes in the atmosphere and between the Earth's surface and atmosphere.	Cloud Particle Phase	Cloud Top Pressure	Cloud Water Path
Latent Heat Flux Latent heat flux is the amount of heat that is transferred across a surface of unit area in a unit of time. Also refers to latent and sensible heat fluxes in the atmosphere and between the Earth's surface and atmosphere.	Project: A-Train AIRS HIRSL MERRA VLS NCEP QM TOVS TRMM	Ground Heat Flux	Heat Latent Diffusivity
Sensible Heat Flux Sensible heat flux is the amount of heat that is transferred across a surface of unit area in a unit of time. Also refers to latent and sensible heat fluxes in the atmosphere and between the Earth's surface and atmosphere.	Downward Heat Flux	Heat Turbulent Flux	Latent Heat Flux
Surface Shortwave Flux Surface shortwave flux is the amount of shortwave radiation that is transferred across a surface of unit area in a unit of time. Also refers to latent and sensible heat fluxes in the atmosphere and between the Earth's surface and atmosphere.	Heat Defectivity	Latent Heat Flux (positive upward)	Sensible Heat Flux
Precipitation Any or all of the forms of water droplets, whether liquid or solid, that fall from clouds and reach the ground or water.	Sensible Heat Flux (positive upward)	Sensible Heat Flux from Land	Total water diffusivity
	Project: GLOAS MERRA	Surface Shortwave Flux	
	Project: MERRA	Atmospheric Water Vapor	Cloud Ice

4

Mirador portal as OpenSearch client